

8



OIQE

RAW SEQUENCE LISTING

DATE: 10/30/2002

PATENT APPLICATION: US/10/086,814

TIME: 13:57:57

Input Set : A:\61010AB1.ST25.txt

Output Set: N:\CRF4\10302002\J086814.raw

3 <110> APPLICANT: Dragic, Tatjana
 4 Olson, William C.
 6 <120> TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
 8 <130> FILE REFERENCE: 61010-AB-1
 10 <140> CURRENT APPLICATION NUMBER: US 10/086,814
 11 <141> CURRENT FILING DATE: 2002-02-28
 13 <160> NUMBER OF SEQ ID NOS: 38
 15 <170> SOFTWARE: PatentIn version 3.1
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 352
 19 <212> TYPE: PRT
 20 <213> ORGANISM: Homo sapiens
 22 <400> SEQUENCE: 1
 24 Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
 25 1 5 10 15
 28 Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
 29 20 25 30
 32 Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
 33 35 40 45
 36 Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
 37 50 55 60
 40 Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
 41 65 70 75 80
 44 Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
 45 85 90 95
 48 Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
 49 100 105 110
 52 Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
 53 115 120 125
 56 Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
 57 130 135 140
 60 Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
 61 145 150 155 160
 64 Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
 65 165 170 175
 68 Thr Cys Ser Ser His Phe Pro Tyr Ser Gln Tyr Gln Phe Trp Lys Asn
 69 180 185 190
 72 Phe Gln Thr Leu Lys Ile Val Ile Leu Gly Leu Val Leu Pro Leu Leu
 73 195 200 205
 76 Val Met Val Ile Cys Tyr Ser Gly Ile Leu Lys Thr Leu Leu Arg Cys
 77 210 215 220
 80 Arg Asn Glu Lys Lys Arg His Arg Ala Val Arg Leu Ile Phe Thr Ile
 81 225 230 235 240

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/086,814

DATE: 10/30/2002

TIME: 13:57:57

Input Set : A:\61010AB1.ST25.txt

Output Set: N:\CRF4\10302002\J086814.raw

```

84 Met Ile Val Tyr Phe Leu Phe Trp Ala Pro Tyr Asn Ile Val Leu Leu
85           245           250           255
88 Leu Asn Thr Phe Gln Glu Phe Phe Gly Leu Asn Asn Cys Ser Ser Ser
89           260           265           270
92 Asn Arg Leu Asp Gln Ala Met Gln Val Thr Glu Thr Leu Gly Met Thr
93           275           280           285
96 His Cys Cys Ile Asn Pro Ile Ile Tyr Ala Phe Val Gly Glu Lys Phe
97           290           295           300
100 Arg Asn Tyr Leu Leu Val Phe Phe Gln Lys His Ile Ala Lys Arg Phe
101 305           310           315           320
104 Cys Lys Cys Cys Ser Ile Phe Gln Gln Glu Ala Pro Glu Arg Ala Ser
105           325           330           335
108 Ser Val Tyr Thr Arg Ser Thr Gly Glu Gln Glu Ile Ser Val Gly Leu
109           340           345           350
112 <210> SEQ ID NO: 2
113 <211> LENGTH: 1376
114 <212> TYPE: DNA
115 <213> ORGANISM: Homo sapiens
117 <400> SEQUENCE: 2
118 gaattccccc aacagagcca agctctccat ctagtggaca gggaagctag cagcaaacct      60
120 tcccttcact acaaaacttc attgcttggc caaaaagaga gttaattcaa tgtagacatc      120
122 tatgtaggca attaaaaacc tattgatgta taaaacagtt tgcattcatg gagggcaact      180
124 aaatacattc taggacttta taaaagatca ctttttattt atgcacaggg tggaacaaga      240
126 tggattatca agtgtcaagt ccaatctatg acatcaatta ttatacatcg gagccctgcc      300
128 aaaaaatcaa tgtgaagcaa atcgcgagccc gcctcctgcc tccgctctac tcactgggtg      360
130 tcatctttgg ttttgtgggc aacatgctgg tcatcctcat cctgataaac tgcaaaaggc      420
132 tgaagagcat gactgacatc tacctgctca acctggccat ctctgacctg ttttcccttc      480
134 ttactgtccc cttctgggct cactatgctg ccgcccagtg ggactttgga aatacaatgt      540
136 gtcaactctt gacagggctc tattttatag gcttcttctc tggaatcttc ttcatcatcc      600
138 tcctgacaat cgataggtac ctggctgtcg tccatgctgt gtttgcttta aaagccagga      660
140 cggtcacctt tgggggtggg acaagtgtga tcacttgggt ggtggctgtg tttgcgtctc      720
142 tcccaggaat catctttacc agatctcaaa aagaaggctt tcattacacc tgcagctctc      780
144 attttccata cagtcagtat caattctgga agaatttcca gacattaaag atagtcatct      840
146 tggggctggt cctgcogctg cttgtcatgg tcatctgcta ctcggaatc ctaaaaactc      900
148 tgcttcgggtg tcgaaatgag aagaagaggc acagggtgtg gaggttatc ttcaccatca      960
150 tgattgttta tttctcttc tgggctccct acaacattgt ccttctcctg aacaccttcc      1020
152 aggaattctt tggcctgaat aattgcagta gctctaacag gttggaccaa gctatgcagg      1080
154 tgacagagac tcttgggatg acgcactgct gcatcaaccc catcatctat gcctttgtcg      1140
156 gggagaagtt cagaaactac ctcttagtct tcttccaaaa gcacattgcc aaacgcttct      1200
158 gcaaagtctg ttctattttc cagcaagagg ctcccagcgc agcaagctca gtttacaccc      1260
160 gatccactgg ggagcaggaa atatctgttg gcttgtgaca cggactcaag tgggctggtg      1320
162 acccagtcag agttgtgcac atggcttagt tttcatacac agcctgggct ggggggt      1376
165 <210> SEQ ID NO: 3
166 <211> LENGTH: 12
167 <212> TYPE: PRT
168 <213> ORGANISM: Homo sapiens
170 <220> FEATURE:
171 <221> NAME/KEY: MISC_FEATURE
172 <222> LOCATION: (1)..(1)

```

RAW SEQUENCE LISTING

DATE: 10/30/2002

PATENT APPLICATION: US/10/086,814

TIME: 13:57:57

Input Set : A:\61010AB1.ST25.txt

Output Set: N:\CRF4\10302002\J086814.raw

```

173 <223> OTHER INFORMATION: Any amino acid
176 <220> FEATURE:
177 <221> NAME/KEY: MISC_FEATURE
178 <222> LOCATION: (2)..(2)
179 <223> OTHER INFORMATION: Xaa represents from 0 to 9 amino acids, where if there are
more t
180     han 2 amino acids, they have a sequence identical to the sequence
181     set forth in SEQ ID NO: 1 beginning with the Ile at position 9 a
182     nd extending therefrom in the amino terminal direction.
185 <220> FEATURE:
186 <221> NAME/KEY: MOD_RES
187 <222> LOCATION: (3)..(3)
188 <223> OTHER INFORMATION: SULFATATION
191 <220> FEATURE:
192 <221> NAME/KEY: MOD_RES
193 <222> LOCATION: (7)..(7)
194 <223> OTHER INFORMATION: SULFATATION
197 <220> FEATURE:
198 <221> NAME/KEY: MISC_FEATURE
199 <222> LOCATION: (11)..(11)
200 <223> OTHER INFORMATION: Xaa represents from 0 to 14 amino acids, where if there are
more
201     than 2 amino acids, they have a sequence identical to the sequenc
202     e set forth in SEQ ID NO: 1 beginning with the Glu at position 18
203     and extending therefrom in the carboxy terminal direction.
206 <220> FEATURE:
207 <221> NAME/KEY: MISC_FEATURE
208 <222> LOCATION: (12)..(12)
209 <223> OTHER INFORMATION: Any amino acid
212 <400> SEQUENCE: 3
W--> 214 Xaa Xaa Tyr Asp Ile Asn Tyr Tyr Thr Ser Xaa Xaa
215 1           5           10
218 <210> SEQ ID NO: 4
219 <211> LENGTH: 12
220 <212> TYPE: PRT
221 <213> ORGANISM: Homo sapiens
223 <220> FEATURE:
224 <221> NAME/KEY: MISC_FEATURE
225 <222> LOCATION: (2)..(2)
226 <223> OTHER INFORMATION: Xaa represents from 0 to 9 amino acids, where if there are
more t
227     han 2 amino acids, they have a sequence identical to the sequence
228     set forth in SEQ ID NO: 1 beginning with the Ile at position 9 a
229     nd extending therefrom in the amino terminal direction.
232 <220> FEATURE:
233 <221> NAME/KEY: MISC_FEATURE
234 <222> LOCATION: (1)..(1)
235 <223> OTHER INFORMATION: Xaa is any amino acid
238 <220> FEATURE:
239 <221> NAME/KEY: MOD_RES
240 <222> LOCATION: (7)..(8)
241 <223> OTHER INFORMATION: SULFATATION

```

RAW SEQUENCE LISTING

DATE: 10/30/2002

PATENT APPLICATION: US/10/086,814

TIME: 13:57:57

Input Set : A:\61010ABL.ST25.txt

Output Set: N:\CRF4\10302002\J086814.raw

```

244 <220> FEATURE:
245 <221> NAME/KEY: MISC_FEATURE
246 <222> LOCATION: (11)..(11)
247 <223> OTHER INFORMATION: Xaa represents from 0 to 14 amino acids, where if there are
more
248     than 2 amino acids, they have a sequence identical to the sequenc
249     e set forth in SEQ ID NO: 1 beginning with the Glu at position 18
250     and extending therefrom in the carboxy terminal direction.
253 <220> FEATURE:
254 <221> NAME/KEY: MISC_FEATURE
255 <222> LOCATION: (12)..(12)
256 <223> OTHER INFORMATION: Xaa is any amino acid
259 <400> SEQUENCE: 4
W--> 261 Xaa Xaa Tyr Asp Ile Asn Tyr Tyr Thr Ser Xaa Xaa
      262 1           5           10
265 <210> SEQ ID NO: 5
266 <211> LENGTH: 12
267 <212> TYPE: PRT
268 <213> ORGANISM: Homo sapiens
270 <220> FEATURE:
271 <221> NAME/KEY: MISC_FEATURE
272 <222> LOCATION: (1)..(1)
273 <223> OTHER INFORMATION: Xaa is any amino acid
276 <220> FEATURE:
277 <221> NAME/KEY: MISC_FEATURE
278 <222> LOCATION: (2)..(2)
279 <223> OTHER INFORMATION: Xaa represents from 0 to 9 amino acids, where if there are
more t
280     han 2 amino acids, they have a sequence identical to the sequence
281     set forth in SEQ ID NO: 1 beginning with the Ile at position 9 a
282     nd extending therefrom in the amino terminal direction.
285 <220> FEATURE:
286 <221> NAME/KEY: MOD_RES
287 <222> LOCATION: (3)..(3)
288 <223> OTHER INFORMATION: SULFATATION
291 <220> FEATURE:
292 <221> NAME/KEY: MOD_RES
293 <222> LOCATION: (8)..(8)
294 <223> OTHER INFORMATION: SULFATATION
297 <220> FEATURE:
298 <221> NAME/KEY: MISC_FEATURE
299 <222> LOCATION: (11)..(11)
300 <223> OTHER INFORMATION: Xaa represents from 0 to 14 amino acids, where if there are
more
301     than 2 amino acids, they have a sequence identical to the sequenc
302     e set forth in SEQ ID NO: 1 beginning with the Glu at position 18
303     and extending therefrom in the carboxy terminal direction.
306 <220> FEATURE:
307 <221> NAME/KEY: MISC_FEATURE
308 <222> LOCATION: (12)..(12)
309 <223> OTHER INFORMATION: Xaa is any amino acid
312 <400> SEQUENCE: 5

```

RAW SEQUENCE LISTING

DATE: 10/30/2002

PATENT APPLICATION: US/10/086,814

TIME: 13:57:57

Input Set : A:\61010AB1.ST25.txt

Output Set: N:\CRF4\10302002\J086814.raw

W--> 314 Xaa Xaa Tyr Asp Ile Asn Tyr Tyr Thr Ser Xaa Xaa

315 1 5 10

318 <210> SEQ ID NO: 6

319 <211> LENGTH: 12

320 <212> TYPE: PRT

321 <213> ORGANISM: Homo sapiens

323 <220> FEATURE:

324 <221> NAME/KEY: MISC_FEATURE

325 <222> LOCATION: (1)..(1)

326 <223> OTHER INFORMATION: Xaa is any amino acid

329 <220> FEATURE:

330 <221> NAME/KEY: MOD_RES

331 <222> LOCATION: (1)..(1)

332 <223> OTHER INFORMATION: ACETYLTATION

335 <220> FEATURE:

336 <221> NAME/KEY: MISC_FEATURE

337 <222> LOCATION: (2)..(2)

338 <223> OTHER INFORMATION: Xaa represents from 0 to 9 amino acids, where if there are
more t

339 han 2 amino acids, they have a sequence identical to the sequence

340 set forth in SEQ ID NO: 1 beginning with the Ile at position 9 a

341 nd extending therefrom in the amino terminal direction.

344 <220> FEATURE:

345 <221> NAME/KEY: MOD_RES

346 <222> LOCATION: (3)..(3)

347 <223> OTHER INFORMATION: SULFATATION

350 <220> FEATURE:

351 <221> NAME/KEY: MOD_RES

352 <222> LOCATION: (7)..(7)

353 <223> OTHER INFORMATION: SULFATATION

356 <220> FEATURE:

357 <221> NAME/KEY: MISC_FEATURE

358 <222> LOCATION: (11)..(11)

359 <223> OTHER INFORMATION: Xaa represents from 0 to 14 amino acids, where if there are
more

360 than 2 amino acids, they have a sequence identical to the sequenc

361 e set forth in SEQ ID NO: 1 beginning with the Glu at position 18

362 and extending therefrom in the carboxy terminal direction.

365 <220> FEATURE:

366 <221> NAME/KEY: MISC_FEATURE

367 <222> LOCATION: (12)..(12)

368 <223> OTHER INFORMATION: Xaa is any amino acid

371 <400> SEQUENCE: 6

W--> 373 Xaa Xaa Tyr Asp Ile Asn Tyr Tyr Thr Ser Xaa Xaa

374 1 5 10

377 <210> SEQ ID NO: 7

378 <211> LENGTH: 12

379 <212> TYPE: PRT

380 <213> ORGANISM: Homo sapiens

382 <220> FEATURE:

383 <221> NAME/KEY: MISC_FEATURE

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/086,814

DATE: 10/30/2002
TIME: 13:57:58

Input Set : A:\61010AB1.ST25.txt
Output Set: N:\CRF4\10302002\J086814.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; Xaa Pos. 1,2,11,12
Seq#:4; Xaa Pos. 1,2,11,12
Seq#:5; Xaa Pos. 1,2,11,12
Seq#:6; Xaa Pos. 1,2,11,12
Seq#:7; Xaa Pos. 1,2,11,12
Seq#:8; Xaa Pos. 1,2,11,12
Seq#:9; Xaa Pos. 1,2,11,12
Seq#:10; Xaa Pos. 1,2,11,12
Seq#:11; Xaa Pos. 1,2,11,12
Seq#:12; Xaa Pos. 1,2,11,12
Seq#:13; Xaa Pos. 1,2,11,12
Seq#:14; Xaa Pos. 1,2,11,12
Seq#:15; Xaa Pos. 1,2,11,12
Seq#:16; Xaa Pos. 1,2,11,12
Seq#:17; Xaa Pos. 1,2,11,12
Seq#:18; Xaa Pos. 1,2,11,12
Seq#:19; Xaa Pos. 1,2,11,12
Seq#:20; Xaa Pos. 1,2,11,12
Seq#:21; Xaa Pos. 1,2,11,12
Seq#:22; Xaa Pos. 1,2,11,12
Seq#:23; Xaa Pos. 1,2,11,12
Seq#:24; Xaa Pos. 1,2,11,12
Seq#:25; Xaa Pos. 1,2,11,12
Seq#:26; Xaa Pos. 1,2,11,12

VERIFICATION SUMMARY

DATE: 10/30/2002

PATENT APPLICATION: US/10/086,814

TIME: 13:57:58

Input Set : A:\61010AB1.ST25.txt

Output Set: N:\CRF4\10302002\J086814.raw

L:214 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:261 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0
L:314 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:373 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
L:426 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0
L:485 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0
L:544 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0
L:597 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0
L:656 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:721 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
L:780 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:845 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0
L:898 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0
L:945 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0
L:998 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:0
L:1057 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:0
L:1110 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0
L:1169 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0
L:1228 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:0
L:1281 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0
L:1340 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:0
L:1405 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:0
L:1464 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:0
L:1529 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:0